Independent market research and competitive analysis of next-generation business and technology solutions for service providers and vendors

HEAVY READING
WHITE PAPER

Beyond the Human Threshold:

5G Policy Control and Charging Strategies

A Heavy Reading white paper produced for STL

AUTHOR: JIM HODGES, RESEARCH DIRECTOR, HEAVY READING

INTRODUCTION

As 5G network deployments continue to ramp up, communications service providers (CSPs) continue to look for new ways to extract value from their core networks and radio access networks (RANs). A vital element of this exercise will be how to achieve further service differentiation from previous mobile generations such as 4G.

While this process will be gradual, the 5G core (5GC), which adopts a services-based architecture (SBA) that supports application programming interface (API) exposure, will drive the next wave of mobile service innovation. Two key components critical to the process of monetizing 5G innovative services are policy control and charging. While both have always been important in mobile networks, they will play an even greater role in 5G due to the additional policy and charging requirements of these complex 5G services.

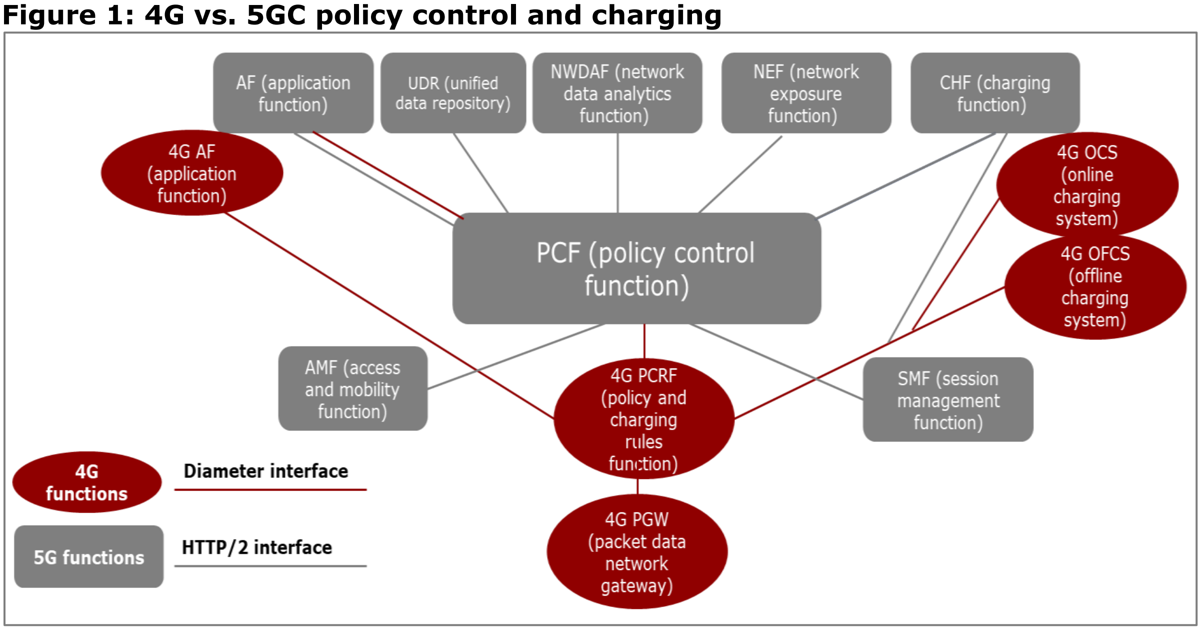
This white paper documents the foundational role that policy control and charging play in the delivery and monetization of complex high value 5G use cases.

THE EVOLUTION OF POLICY CONTROL AND CHARGING

Since 5G standards define a new 5GC network based on cloud native architecture, functions such as policy control and charging that reside in the core differ significantly from 4G networks. In 4G core networks, policy and charging capabilities are supported by the policy and charging rules function (PCRF), application function (AF), online charging system (OCS), and offline charging system (OFCS), which utilize Diameter interfaces.

In contrast, in the 5GC realm, policy and charging are supported by the policy control function (PCF), AF, and charging function (CHF). These functions utilize HTTP/2 protocol-based control plane interfaces referred to as service-based interfaces (SBIs). **Figure 1** illustrates how these functions are mapped between 4G and 5G networks.

Figure 1: 4G vs. 5GC policy control and charging



Source: Heavy Reading

The mapping of 4G policy and charging functions to 5G policy and charging functions is similar by design to facilitate 4G/5G backward compatible service interworking. However, it is important to note that these 5G functions are now supported in a fully cloud native environment.

Therefore, while these nodes appear similar, as documented below, 5G policy and charging functions are more advanced to support complex capabilities such as slice-based services and URLLC (ultra-reliable low latency communications) cloud native services.

4G PCRF vs. 5G PCF

To meet these rigid service demands, the 5G PCF utilizes an HTTP/2-based interface between the PCF and network exposure function (NEF), network data and analytics function (NWDAF), and CHF. Like the CHF, the AF and the HTTP/s interfaces also support Diameter-based interfaces to simplify 4G service delivery.

This shift to HTTP/2 is significant since it aligns with the cloud domain and enables 5G services to be run not only in a private CSP cloud but also in a public cloud. In addition, the new protocol is designed to enable the exposure of APIs via the NEF—a new capability not supported in the 4G world.

There are two considerations here. The first is the ability to support and enforce strict policy-based quality of service (QoS) tolerances for these API-based services. Another important difference between PCF and PCRF is the fact that PCF administers policies for 5GC and radio network (e.g., gNB) functions, which is vital to support policies for network slices. In contrast, the PCRF supports policy enforcement for core network nodes only.

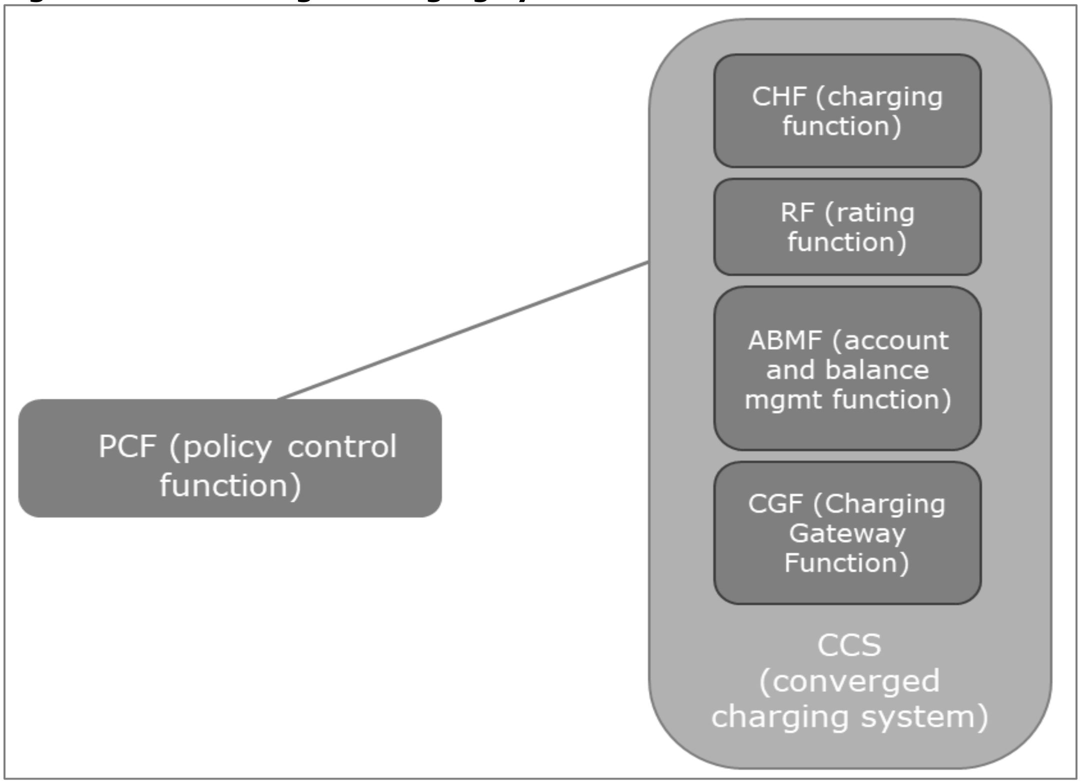
An additional consideration relates to the software origins of the API themselves. In the 5G realm, there is also a much greater focus on the integration of third-party APIs vital to fostering the adoption of an ecosystem-based service delivery model that is foundational to the successful monetization of 5G services. Since these APIs may also run in a slice-based environment to meet strict latency budgets, a shift from Diameter is also warranted.

4G OCS vs. 5G CHF

The evolution to a cloud native architecture is equally profound from a charging perspective. As shown below in **Figure 2**, the new 5G cloud native enabled converged charging system (CCS) is defined in standards that include the CHF to support charging of complex 5G services. They also include a rating function (RF) to manage the rating of stateless microservices, an account and balance management function (ABMF) to maintain subscriber account balances, and a charging gateway function (CGF) to manage the format for billing record generation in a virtualized environment.

The 5G CCS introduces new control plane interfaces swapping out Diameter protocol-based interfaces for HTTP/2-based SBI.

Figure 2: 5G converged charging system



Source: Heavy Reading

Like the policy function, these converged charging functions will accelerate the 5G services monetization curve. They can support existing 4G services with the performance and flexibility capabilities to support advanced cloud-based 5G uses cases when working in tandem with the PCF.

5G POLICY AND CHARGING SYSTEM DESIGN ATTRIBUTES

To achieve the promise of the technology innovation documented in the previous section, CSPs' 5G policy and execution strategies and the policy and charging systems they deploy must support the following foundational design attributes:

- Scale
- Programmability
- Standardization
- Automation

Scale

The ability of an architecture to scale has always represented a vital network metric. However, in many respects, 5G takes the concept of scalability to an entirely new level. Over the next 8–10 years, 5G technology and Internet of Things (IoT) devices will be embedded in autonomous vehicles and industrial sites to support both public and private network implementations.

A related consideration here is not simply the number of devices and nodes that must be supported, but also the exponential increase in the amount of performance-related data these devices will generate. As a result, policy and converged charging systems will need to support massive scale requirements to enable a broad range of high-value use cases, as documented in the next section of this white paper.

Programmability

One approach to achieve massive scale is to support 5G service invocation anywhere in the network: at the edge and in both private and public clouds.

To support this model, programmability, the ability to utilize software to tailor service performance based on where service execution takes place is critical. This requirement, in turn, requires that policy and charging systems become more programmable as well to respond in real time to this changing services landscape.

Standardization

Although standards have always been important for ensuring servicer provider interworking, since policy and charging must not be supported in a multicloud environment, the overall rigor of standardization becomes even more important to ensure seamless interworking. In this respect, the development and support of comprehensive standards are vital elements in monetizing programmability in specific use cases.

Automation

The final interrelated attribute is automation. While 5G network design and standards do not specifically mandate the support of automated policy control and charging, as 5G services scale, Heavy Reading believes transitioning to automated systems is inevitable.

The key driver in this transition is related to the evolution of network design principles. 2G, 3G, and to some extent 4G mobile networks are designed to conform to human performance tolerances. The principle here is that network performance latency budgets are based on human tolerances.

In the 2G world, this resulted defining the optimal delay for applying ringing after dialing. This was an important step since applying ringing too early or too late would cause voice subscribers to abandon calls, resulting in additional call attempts and ultimately degraded customer quality metrics.

However, in the 5G world, many services such as autonomous vehicles are machine-driven and not subject to human performance tolerance limitations. In this new world order, automation is critical to meet higher machine-to-machine (M2M) performance expectations.

Stated differently, to be successful, the 5G era policy and charging systems must be capable of evolving beyond the human threshold to support M2M performance thresholds.

5G MONETIZATION USE CASES

As documented, 5G service invocation will have major policy and charging implications. This section documents the implications on a more granular level by examining the enhanced role that policy control and charging play in the following 5G use cases:

- Everything as a service (XaaS) – mobile virtual network operator (MVNO)
- 5G slice-based enterprise use case
- 5G private network – IoT manufacturing integration

Everything as a service (XaaS) – MVNO

One of the distinct advantages of migrating services to the cloud is that it enables CSPs to gain access to new business opportunities. One example is the ability to sell XaaS to enterprise customers.

This is important because XaaS leverages scale and zero-touch automation, enabling CSPs to deliver advanced, scalable services to consumers, enterprise customers, or even MNVOs that are hosted on the CSP network. XaaS is gaining market momentum since previous mobile network generations did not have the capability and network performance speeds, programmability, and openness to support the foundational requirements of XaaS.

The use case illustrated in **Figure 3** below describes a CSP that supports the service delivery and billing for a 5G MNVO customer. In this case, the MVNO customers are provided access to a cloud-based self-service portal where they can provision and activate new service capabilities in real time.

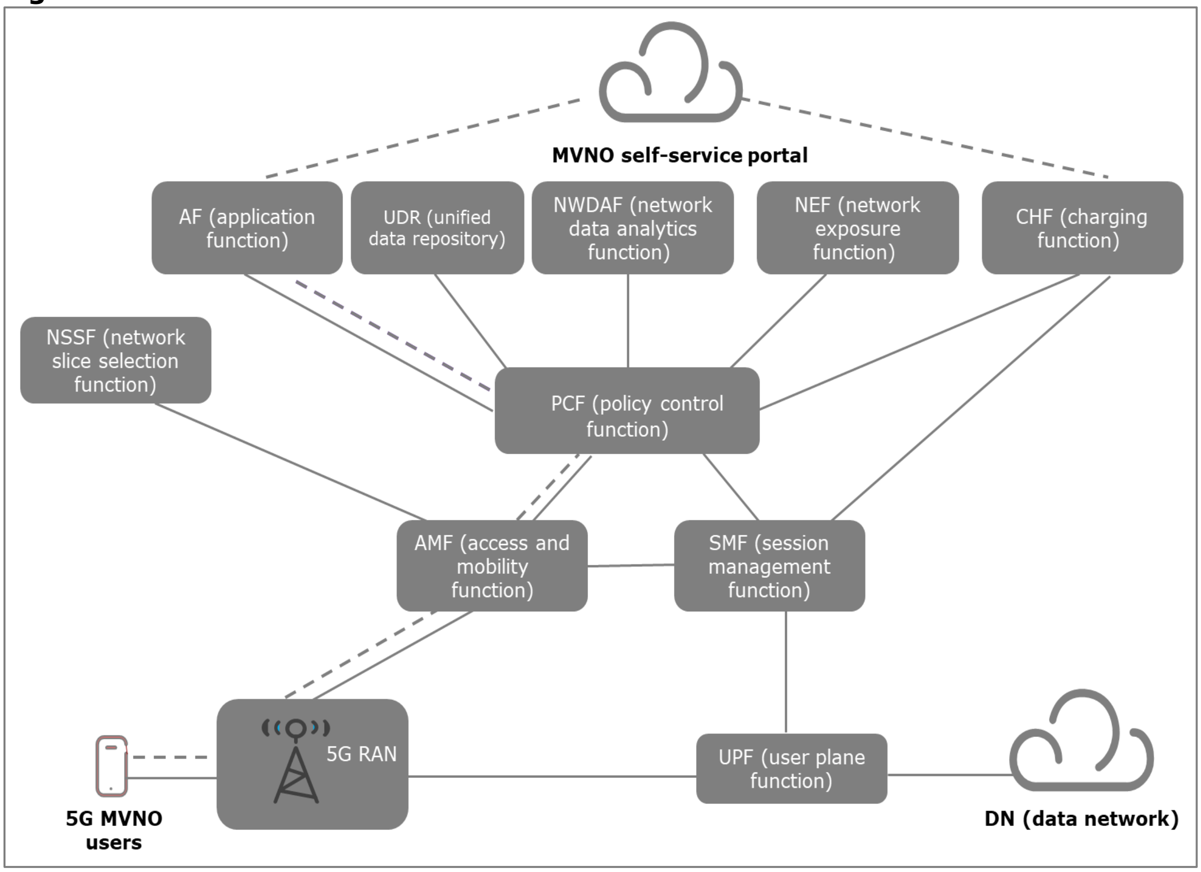
This has all the policy and charging requirements of normal 5G customers (e.g., data quota management). It also includes the additional functions associated with monitoring and recording transactions so the CSP can bill the MNVO for managing subscriber additions and deletes.

This model empowers the MVNO to streamline service provisioning while providing the end user with an enhanced user experience by granting these subscribers a greater measure of control and “feel” for their user experience. The above use case also illustrates how CSPs are evolving to the commercialization of the cloud through the adoption of new business value-add models.

One example of a new model is the emergence of the business-to-business-to-X (B2B2x). Effectively, this approach extends the traditional business-to-business (B2B) model by allowing the second "B" in the value chain the opportunity to take the service it has purchased and add value to the end user.

In this use case, the CSP sells network access to the MVNO (the second "B"), which in turn utilizes the self-service portal to deliver the end-user customer (the "X") a more flexible and richer user experience.

Figure 3: 5G MNVO XaaS use case



Source: Heavy Reading

For XaaS to be monetizable in any model (B2B, B2B2x, or business-to-consumer [B2C]), there are unique policy and charging platform requirements.

Scalability

As noted, scalability is a fundamental attribute of XaaS that enables CSPs to drive new revenue streams while at the same time lowering the total cost of ownership (TCO) to enhance profitability. To be successful here, CSPs' 5G policy and charging systems must be able to manage the increased workloads associated with XaaS.

Programmability

One of the main benefits of XaaS is that it represents a blank template from a service delivery perspective. As 5G services evolve, CSPs will be able to offer new XaaS-based services aligned with industry adoption trends. CSPs have this flexibility since XaaS is fully programmable and can therefore be tailored to meet man or machine service needs without extensive programming requirements.

This means that policy and charging systems must also be fully programmable. They must have the flexibility to support any charging model, including event-based charging and quota-based charging, based on XaaS market requirements.

Effectively, this means programmable policy and charging systems will evolve to support charging as a service (ChaaS) for any service, including nontraditional CSP services. CSPs will have the opportunity to enter new market segments as a trusted ecosystem partner.

Standardization

One of the key standardization considerations for XaaS services is API-related. As documented, 5G's API exposure model is highly complementary to XaaS service delivery.

In this model, APIs must be standardized to enable CSPs to seamlessly offer any service, whether the XaaS service runs in the telco cloud or public cloud. Consequently, policy and charging systems must be fully multi-standard compliant to ensure they seamlessly apply policy and charge for API-related events.

Automation

The addition of the self-service portal in this use case delivers considerable value, but it also drives additional policy and charging requirements. This is because the portal can be viewed as representing a flexible and extensible library of services. With this library, each service will have a unique set of policy rules and charging requirements. To manage this policy, systems must support automated rule sets that can turn up or tear down specific services based on network performance or even security demands.

Moreover, since this portal will be constantly evolving to support new services, automation will be a valuable tool in allowing policy and charging systems to adjust to new service requirements in real time.

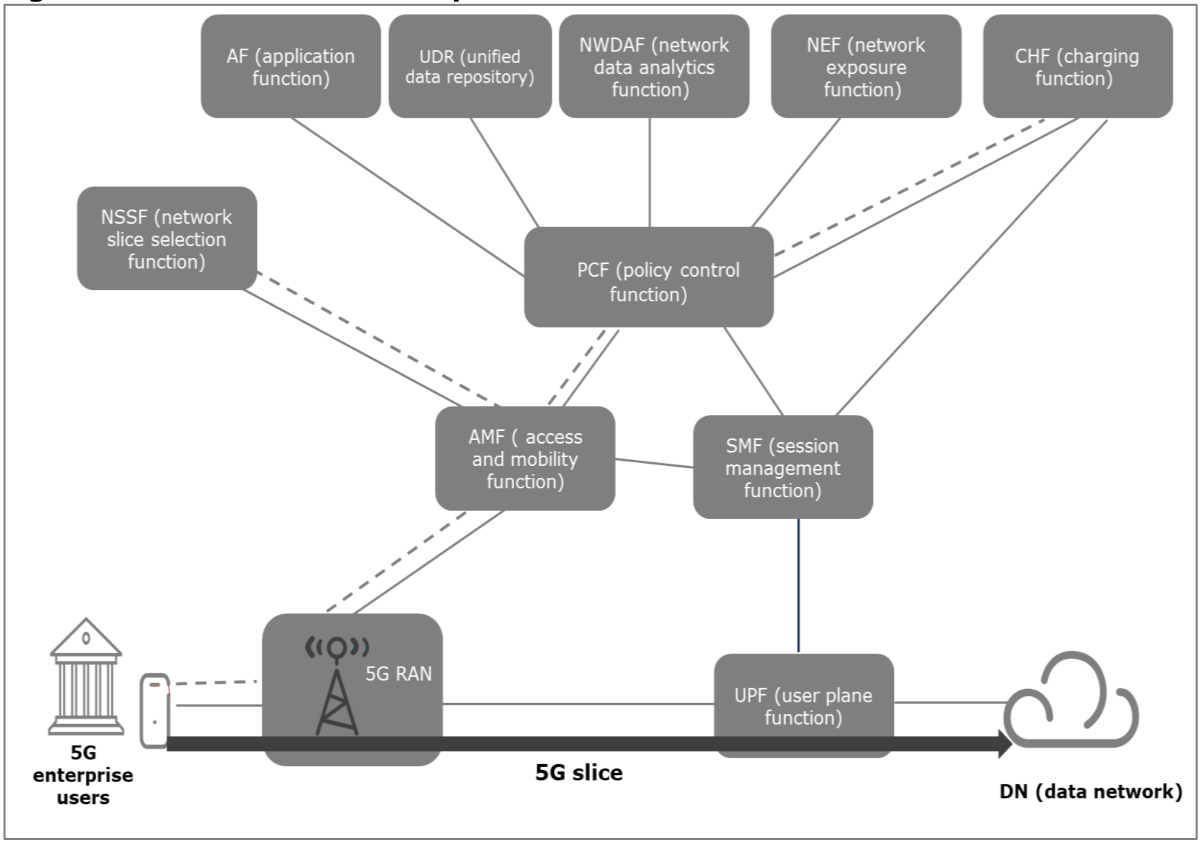
5G slice-based enterprise use case

Enterprises are expected to be heavy users of slice-based services. This is because slicing can be tailored to support the unique performance and scale demands of individual enterprises.

This use case is based on a medium-sized financial consulting enterprise that relies on network slicing to enable the processing and billing of financial analyst engagements with end-user customers.

This business demands agility and performance, so it utilizes slice-based connectivity to provide benchmark performance and secure transactions, as illustrated in **Figure 4**. Therefore, it has strict service-level agreement (SLA) and policy and billing requirements.

Figure 4: 5G slice-based enterprise use case



Source: Heavy Reading

As with the other use cases presented in this white paper, monetization success will hinge on policy and charging platform design strategies that embody the following design attributes.

Scalability

In this use case, scalability is critical due to end-user customer expectations of unfettered access to financial analysts and rapid transaction execution.

Programmability

Policy control and charging enforcement are critical for administering enterprise services. In this scenario, programmable systems are crucial for ensuring specific enterprise users in specific locations for specific services are provided the slice-based QoS they have paid for. Meanwhile, other enterprise customers who are not allowed access based on their subscriber profile will be blocked or will have their sessions terminated.

There is another wrinkle in the enterprise model that necessitates CSPs deploy programmable policies and charging systems. For example, in this use case, the financial organization in a remote location needs to offer the same level of support for their financial analysts as their major enterprise offices. Since the CSP may not have sufficient radio resources in this location, it may partner with a "neutral host" third-party company that has radio or transport resources in this resource region that the CSP can integrate into its enterprise slicing service. While the enterprise experiences a single look and feel from a single CSP, this same CSP must be able to pull in RAN and analytics data into policy and charging systems to make the slice service truly seamless.

Standardization

The value of network slicing in this use case is that it enables enterprise customers to have the feel of a bespoke virtual network with guaranteed performance tolerances on a single physical network.

This is a game-changer for enterprises since this level of performance could not be guaranteed on previous-generation mobile networks. For this capability to be delivered, only open standardized interfaces and protocols can be supported to meet these end-to-end performance and online charging demands.

Automation

Enterprises such as financial entities that manage critical big data will also fuel the adoption of artificial intelligence (AI)-based applications to lower costs and process real time transactions. This AI adoption curve represents another factor driving policy and charging systems to integrate automation and algorithm policy capabilities.

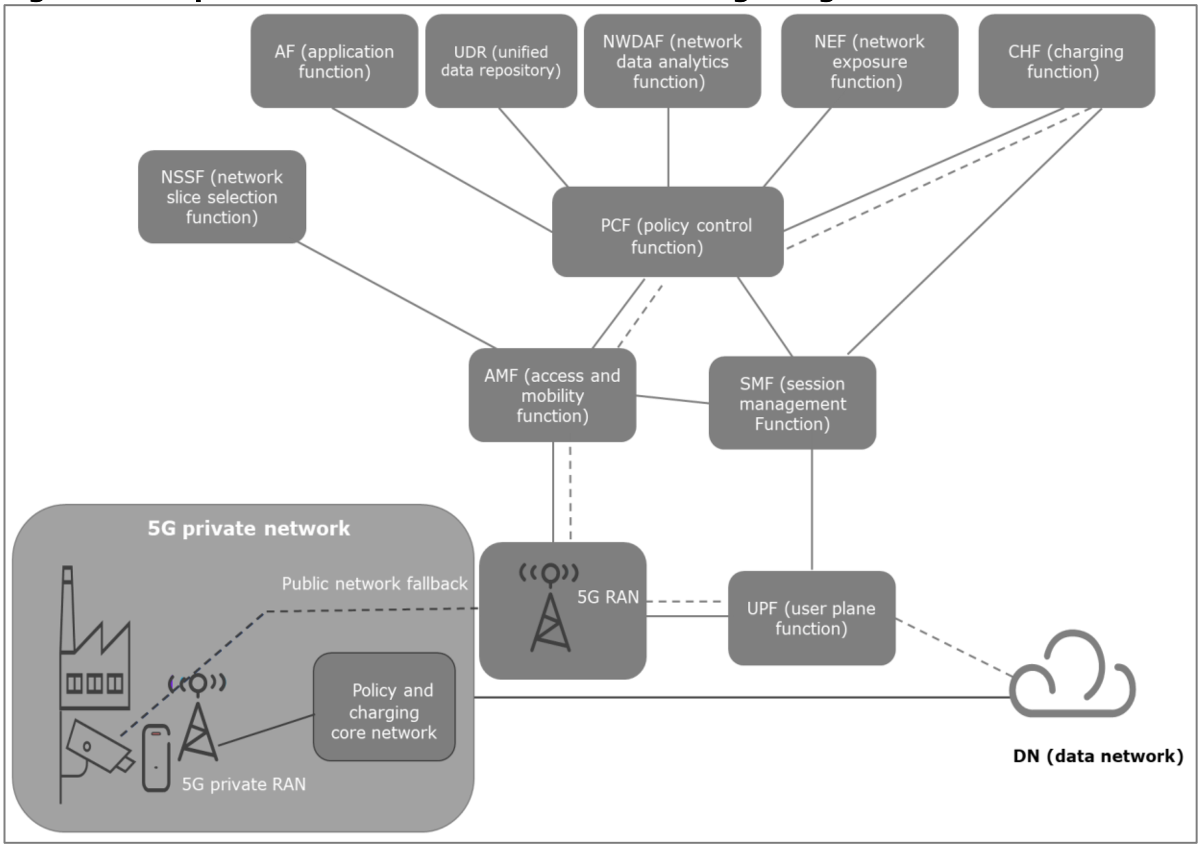
5G private network – IoT manufacturing integration

This final use case falls into the beyond human threshold realm and illustrates a 5G private network deployment for a smart manufacturer that relies on IoT and robotic devices in the manufacturing, surveillance, and quality control processes. This deployment exemplifies the Industry 5.0 model in which manufacturing shifts to an intelligent data-driven operational model that leverages the power of IoT and robotics.

As captured in **Figure 5** below, the value of the private network in this use case is that it provides the manufacturer a greater measure of control and hence security in the manufacturing process. It also enables the management of complex automated supply chain processes that introduce additional policy and charging requirements.

One consideration fueling the additional requirements documented in **Figure 5** is the usage of telco network fallback for the support of critical devices such as production and surveillance cameras. While this approach provides greatly needed survivability if onsite private network policy capabilities malfunction, it requires that policy and potentially charging functions in both private and telco network domains work seamlessly when failover occurs.

Figure 5: 5G private network – IoT manufacturing integration use case



Source: Heavy Reading

This use case also possesses unique policy and charging platform attributes that will be compulsory for effective service monetization.

Scalability

Since this factory is heavily automated with devices that are supported by onsite policy and charging, scale is even a consideration in this use case. This ability to support local onsite processing is valuable because it foregoes the requirement to backhaul transactions to the CSP's centralized charging function.

Programmability

Private networks are not new. However, there has been a real increase in the momentum of private networks, in large part due to the rollout of 5G networks. Essentially, 5G now provides the performance metrics critical to overcoming the previous performance limitations of earlier private network deployments.

A second and arguably more important factor in private network momentum is the programmable aspect of private networks. Since 5G implements microservices that can be implemented onsite, it is possible to reuse and reprogram services in machine time as necessary. To support this, policy and charging systems must also be highly programmable to avoid the creation of service-impacting billing and policy enforcement bottlenecks.

Standardization

The ability of policy and charging systems to support open standardized interfaces and message format in this use case is also vital. Since this private network supports network failover, any policy and charging capabilities deployed in the factory must be fully non-proprietary and fully configurable with the CSP's network capabilities to ensure seamless failover.

Automation

Like the slicing use case, this private network will lean heavily on automated policy and charging tools. In this use case, automation is critical because of the number and speed of the machine devices that will have to be managed.

CONCLUSION

CSPs are now undertaking a major 5G-fueled network transformation that will drive them to reevaluate how they deliver services and charge for them.

One of the hallmarks of this journey is a focus on new service delivery models and use cases that will exceed human performance thresholds. And as captured in this white paper, policy and charging systems will need to respond to this new complex world order with a greater focus on scalable, programmable, standards-based, and automated solutions that can manage the stringent requirements of both 4G and 5G networks.